

**Remarque importante 6.42**

Attention à ne pas confondre  $UU^T$  et  $U^TU$  !

**6.6 Matrices orthogonales****Théorème 6.43**

Soit  $U \in \mathcal{M}_{m,n}(\mathbb{R})$  une matrice de taille  $m \times n$  dont les colonnes sont orthonormées. Si  $\vec{v}, \vec{w} \in \mathbb{R}^n$ , alors :

1.  $\|U\vec{v}\| = \|\vec{v}\|$ ,
2.  $(U\vec{v}) \cdot (U\vec{w}) = \vec{v} \cdot \vec{w}$ ,
3.  $(U\vec{v}) \cdot (U\vec{w}) = 0 \iff \vec{v} \cdot \vec{w} = 0$ .

Dans ce cas, l'application linéaire  $T : \begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ \vec{v} & \longmapsto & U\vec{v} \end{array}$  conserve les longueurs et l'orthogonalité et elle est injective.

*Démonstration.* Comme les colonnes de  $U$  sont orthonormées, on sait que

$$U^TU = I_n.$$

1. Pour la norme, on calcule

$$\begin{aligned} \|U\vec{v}\|^2 &= (U\vec{v}) \cdot (U\vec{v}) = (U\vec{v})^T(U\vec{v}) \\ &= \vec{v}^T U^T U \vec{v} = \vec{v}^T I_n \vec{v} = \vec{v}^T \vec{v} = \|\vec{v}\|^2. \end{aligned}$$

En prenant la racine, on obtient  $\|U\vec{v}\| = \|\vec{v}\|$ .

On a donc que  $T$  est injective, puisque  $U\vec{v} = \vec{0} \iff \|U\vec{v}\| = 0 \iff \|\vec{v}\| = 0 \iff \vec{v} = \vec{0}$ .

2. Pour le produit scalaire,

$$\begin{aligned} (U\vec{v}) \cdot (U\vec{w}) &= (U\vec{v})^T(U\vec{w}) \\ &= \vec{v}^T U^T U \vec{w} \\ &= \vec{v}^T I_n \vec{w} \\ &= \vec{v}^T \vec{w} = \vec{v} \cdot \vec{w}. \end{aligned}$$

3. Le point 2 donne directement

$$(U\vec{v}) \cdot (U\vec{w}) = \vec{v} \cdot \vec{w},$$

d'où

$$(U\vec{v}) \cdot (U\vec{w}) = 0 \iff \vec{v} \cdot \vec{w} = 0.$$

□

**Exemple.** Soit

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix}$$

Les colonnes

$$\vec{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \quad \text{et} \quad \vec{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

sont orthonormées.

Prenons  $\vec{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$  et  $\vec{w} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$  dans  $\mathbb{R}^2$ .

On calcule d'abord :

$$U\vec{v} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} \\ 1 \end{bmatrix}$$

$$U\vec{w} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ -3 \end{bmatrix}$$

**Conservation de la norme.**

$$\|\vec{v}\| = \sqrt{3^2 + 1^2} = \sqrt{10}$$

$$\|U\vec{v}\| = \sqrt{\left(\frac{3}{\sqrt{2}}\right)^2 + \left(\frac{3}{\sqrt{2}}\right)^2 + 1^2} = \sqrt{\frac{9}{2} + \frac{9}{2} + 1} = \sqrt{10}$$

Donc  $\|U\vec{v}\| = \|\vec{v}\|$ .

De même :

$$\|\vec{w}\| = \sqrt{1^2 + (-3)^2} = \sqrt{10}$$

$$\|U\vec{w}\| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 + (-3)^2} = \sqrt{\frac{1}{2} + \frac{1}{2} + 9} = \sqrt{10}$$

Donc  $\|U\vec{w}\| = \|\vec{w}\|$ .

**Conservation du produit scalaire.**

$$\vec{v} \cdot \vec{w} = 3 \cdot 1 + 1 \cdot (-3) = 3 - 3 = 0$$

$$(U\vec{v}) \cdot (U\vec{w}) = \frac{3}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{3}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + 1 \cdot (-3) = \frac{3}{2} + \frac{3}{2} - 3 = 0$$

Donc  $(U\vec{v}) \cdot (U\vec{w}) = \vec{v} \cdot \vec{w}$ .

L'application linéaire  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  définie par  $T(\vec{x}) = U\vec{x}$  conserve donc les longueurs et l'orthogonalité. Les colonnes de  $U$  sont indépendantes, donc  $T$  est injective.

**Définition 6.44**

*Matrice orthogonale*

On dit qu'une matrice carrée  $A \in \mathcal{M}_n(\mathbb{R})$  est orthogonale si  $A$  est inversible et  $A^{-1} = A^T$ .

**Remarque 6.6.0.45.** Comme par définition, nous avons  $AA^T = \mathbb{I}_n$ , les colonnes de  $A$  sont orthonormées, d'après le théorème 6.41, et forment donc une base orthonormale de  $\mathbb{R}^n$ .

C'est aussi le cas des lignes de  $A$ , puisque  $A^T A = \mathbb{I}_n$ .

**Exemples.** 1. **Matrice identité.** Pour tout  $n \geq 1$ , la matrice identité  $\mathbb{I}_n$  est orthogonale car  $\mathbb{I}_n^T \mathbb{I}_n = \mathbb{I}_n$ .

2. **Matrice de rotation dans  $\mathbb{R}^2$ .** Pour tout angle  $\theta \in \mathbb{R}$ , la matrice

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

est orthogonale. En effet :

$$R_\theta^T R_\theta = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Par exemple, pour  $\theta = \frac{\pi}{4}$  :

$$R_{\pi/4} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

3. **Matrice de réflexion.** La matrice

$$S = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

représente une réflexion par rapport à l'axe des  $x$  et est orthogonale car  $S^T S = S^2 = \mathbb{I}_2$ .

Plus généralement, pour  $a, b \in \mathbb{R}$  avec  $a^2 + b^2 = 1$ , la matrice de réflexion par rapport à la droite engendrée par  $\begin{bmatrix} a \\ b \end{bmatrix}$  est :

$$H = \begin{bmatrix} a^2 - b^2 & 2ab \\ 2ab & b^2 - a^2 \end{bmatrix}$$

qui est orthogonale.

4. **Matrice de permutation.** La matrice

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

qui permute circulairement les coordonnées est orthogonale car  $P^T P = \mathbb{I}_3$ .

5. **Matrices de Householder.** Pour tout vecteur unitaire  $\vec{u} \in \mathbb{R}^n$  (avec  $\|\vec{u}\| = 1$ ), la matrice

$$H = \mathbb{I}_n - 2\vec{u}\vec{u}^T$$

est orthogonale. Par exemple, pour  $\vec{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$  dans  $\mathbb{R}^3$  :

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} [1 \ 0 \ 0] = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

6. **Matrice diagonale à coefficients  $\pm 1$ .** La matrice

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

est orthogonale car  $D^T D = D^2 = \mathbb{I}_3$ .

7. **Produit de matrices orthogonales.** Si  $A$  et  $B$  sont des matrices orthogonales de même taille, alors  $AB$  est aussi orthogonale car :

$$(AB)^T(AB) = B^T A^T AB = B^T I_n B = B^T B = \mathbb{I}_n$$

## 6.7 Le procédé d'orthogonalisation de Gram-Schmidt

Le procédé de Gram-Schmidt permet de transformer une base quelconque en une base orthogonale (ou orthonormée).

### Méthode 6.46

### Procédé de Gram-Schmidt

Soit  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$  un ensemble linéairement indépendant dans  $\mathbb{R}^n$ .

On construit une base orthogonale  $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$  de  $\text{Vect}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k)$  comme suit :

$$\begin{aligned} \vec{u}_1 &= \vec{v}_1 \\ \vec{u}_2 &= \vec{v}_2 - \text{proj}_{\vec{u}_1} \vec{v}_2 \\ \vec{u}_3 &= \vec{v}_3 - \text{proj}_{\vec{u}_1} \vec{v}_3 - \text{proj}_{\vec{u}_2} \vec{v}_3 \\ &\vdots \\ \vec{u}_k &= \vec{v}_k - \sum_{i=1}^{k-1} \text{proj}_{\vec{u}_i} \vec{v}_k \end{aligned}$$

Pour obtenir une base orthonormée, on normalise :  $\vec{q}_i = \frac{\vec{u}_i}{\|\vec{u}_i\|}$ .

**Remarques 6.6.0.47. Pourquoi le procédé fonctionne-t-il ?** À chaque étape  $j$ , le vecteur  $\vec{u}_j$  est construit en retirant de  $\vec{v}_j$  toutes ses composantes dans les directions  $\vec{u}_1, \dots, \vec{u}_{j-1}$ .

Ce qui reste est donc orthogonal à tous les vecteurs précédents. De plus, comme  $\vec{v}_j \notin \text{Vect}(\vec{v}_1, \dots, \vec{v}_{j-1})$  (par indépendance linéaire), le vecteur  $\vec{u}_j$  est nécessairement non nul.

**Espace engendré** Pour tout  $j \in \{1, \dots, k\}$ , on a :

$$\text{Vect}(\vec{u}_1, \dots, \vec{u}_j) = \text{Vect}(\vec{v}_1, \dots, \vec{v}_j)$$

$$\text{En particulier, } \text{Vect}(\vec{u}_1, \dots, \vec{u}_k) = \text{Vect}(\vec{v}_1, \dots, \vec{v}_k).$$

**Exemple.** Orthogonalisons la base  $\left( \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right)$  de  $\mathbb{R}^3$ .

$$\text{Étape 1 : } \vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

**Étape 2 :**

$$\text{proj}_{\vec{u}_1} \vec{v}_2 = \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$\vec{u}_2 = \vec{v}_2 - \text{proj}_{\vec{u}_1} \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

**Étape 3 :**

$$\text{proj}_{\vec{u}_1} \vec{v}_3 = \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$\text{proj}_{\vec{u}_2} \vec{v}_3 = \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2 = \frac{1/2}{3/2} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

$$\vec{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

On obtient la base orthogonale  $\left( \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2/3 \\ -2/3 \\ 2/3 \end{bmatrix} \right)$ .

**Normalisation :** Pour obtenir une base orthonormée, calculons les normes et normalisons :

$$\|\vec{u}_1\| = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}$$

$$\|\vec{u}_2\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + 1^2} = \sqrt{\frac{3}{2}} = \frac{\sqrt{6}}{2}$$

$$\|\vec{u}_3\| = \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2 + \left(\frac{2}{3}\right)^2} = \frac{2\sqrt{3}}{3}$$

La base orthonormée est donc :

$$\left( \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \frac{\sqrt{6}}{6} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \frac{\sqrt{3}}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right)$$

**Exemple.** Orthogonalisons l'ensemble  $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$  de  $\mathbb{R}^4$  par le procédé de Gram-Schmidt.

On note

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

**Étape 1 :**

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

**Étape 2 :**

$$\text{proj}_{\vec{u}_1} \vec{v}_2 = \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.$$

On calcule

$$\vec{v}_2 \cdot \vec{u}_1 = 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 = 1, \quad \vec{u}_1 \cdot \vec{u}_1 = 1^2 + 1^2 = 2,$$

d'où

$$\text{proj}_{\vec{u}_1} \vec{v}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix}.$$

On pose alors

$$\vec{u}_2 = \vec{v}_2 - \text{proj}_{\vec{u}_1} \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix}.$$

On vérifie que  $\vec{u}_2 \cdot \vec{u}_1 = 0$ .

**Étape 3 :**

$$\text{proj}_{\vec{u}_1} \vec{v}_3 = \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.$$

Comme

$$\vec{v}_3 \cdot \vec{u}_1 = 1 \cdot 1 + 0 \cdot 1 + 0 \cdot 0 + 1 \cdot 0 = 1,$$

on obtient

$$\text{proj}_{\vec{u}_1} \vec{v}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix}.$$

Pour la projection sur  $\vec{u}_2$  :

$$\text{proj}_{\vec{u}_2} \vec{v}_3 = \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2.$$

On calcule

$$\vec{v}_3 \cdot \vec{u}_2 = 1 \cdot \frac{1}{2} + 0 \cdot \left(-\frac{1}{2}\right) + 0 \cdot 1 + 1 \cdot 0 = \frac{1}{2},$$

$$\vec{u}_2 \cdot \vec{u}_2 = \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 1^2 + 0^2 = \frac{1}{4} + \frac{1}{4} + 1 = \frac{3}{2},$$

donc

$$\frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} = \frac{1/2}{3/2} = \frac{1}{3},$$

et

$$\text{proj}_{\vec{u}_2} \vec{v}_3 = \frac{1}{3} \vec{u}_2 = \frac{1}{3} \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/6 \\ -1/6 \\ 1/3 \\ 0 \end{bmatrix}.$$

On pose enfin

$$\begin{aligned} \vec{u}_3 &= \vec{v}_3 - \text{proj}_{\vec{u}_1} \vec{v}_3 - \text{proj}_{\vec{u}_2} \vec{v}_3 \\ &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/6 \\ -1/6 \\ 1/3 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 - \frac{1}{2} - \frac{1}{6} \\ 0 - \frac{1}{2} + \frac{1}{6} \\ 0 - 0 - \frac{1}{3} \\ 1 - 0 - 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \\ 1 \end{bmatrix}. \end{aligned}$$

On vérifie que

$$\vec{u}_1 \cdot \vec{u}_3 = 0 \quad \text{et} \quad \vec{u}_2 \cdot \vec{u}_3 = 0,$$

donc les vecteurs  $\vec{u}_1, \vec{u}_2, \vec{u}_3$  sont deux à deux orthogonaux.

On a ainsi obtenu une base orthogonale de  $\text{Vect}(\vec{v}_1, \vec{v}_2, \vec{v}_3)$  :

$$\mathcal{B}_{\text{ortho}} = \left( \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/3 \\ -1/3 \\ -1/3 \\ 1 \end{bmatrix} \right).$$

Si l'on souhaite une base orthonormée, il suffit de normaliser chacun des vecteurs  $\vec{u}_1, \vec{u}_2, \vec{u}_3$ .

## 6.8 Factorisation QR

La factorisation QR est très utile dans la résolution numérique de systèmes  $A\vec{x} = \vec{b}$ , où  $A$  est une matrice de taille  $m \times n$  avec  $m$  et  $n$  très grands. Elle peut aussi fournir un moyen rapide de calculer des valeurs propres d'une matrice.

### Définition 6.48

### Factorisation QR

Soit  $A$  une matrice  $m \times n$  dont les colonnes sont linéairement indépendantes.

La factorisation QR de  $A$  est :  $A = QR$

où :

- $Q$  est une matrice  $m \times n$  dont les colonnes forment une base orthonormée de  $\text{Im}(A)$
- $R$  est une matrice  $n \times n$  triangulaire supérieure inversible avec des éléments diagonaux strictement positifs

### Théorème 6.49

### Théorème de factorisation QR

Toute matrice  $A$  de taille  $m \times n$  avec  $\text{rang}(A) = n$  admet une factorisation QR unique.

*Démonstration.* On suppose que  $A \in \mathcal{M}_{m,n}(\mathbb{R})$  vérifie  $\text{rang}(A) = n$ . Les  $n$  colonnes de  $A$  sont alors linéairement indépendantes et forment une base  $(\vec{a}_1, \dots, \vec{a}_n)$  de  $\text{Im}(A)$ , le sous-espace vectoriel engendré par les colonnes de  $A$ .

En appliquant le procédé d'orthogonalisation de Gram-Schmidt à cette base, on obtient une base orthogonale  $(\vec{u}_1, \dots, \vec{u}_n)$  de  $\text{Im}(A)$ .

En normalisant chaque vecteur,

$$\vec{q}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|}, \quad \vec{q}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|}, \quad \dots, \quad \vec{q}_n = \frac{\vec{u}_n}{\|\vec{u}_n\|},$$

on obtient une base orthonormale  $(\vec{q}_1, \dots, \vec{q}_n)$  de  $\text{Im}(A)$ .

On pose alors

$$Q = [\vec{q}_1 \quad \vec{q}_2 \quad \dots \quad \vec{q}_n].$$

Par construction, les colonnes de  $Q$  sont orthonormées, donc

$$Q^T Q = I_n.$$

D'autre part, le procédé de Gram-Schmidt montre que chaque colonne  $\vec{a}_j$  de  $A$  s'écrit

comme combinaison linéaire de  $\vec{q}_1, \dots, \vec{q}_j$  seulement. On peut donc écrire

$$\vec{a}_j = r_{1j}\vec{q}_1 + r_{2j}\vec{q}_2 + \dots + r_{jj}\vec{q}_j$$

pour certains coefficients  $r_{ij}$ . En rassemblant ces coefficients dans une matrice  $R \in \mathcal{M}_n(\mathbb{R})$ , on obtient une matrice triangulaire supérieure (car  $r_{ij} = 0$  pour  $i > j$ ), et

$$A = QR.$$

En multipliant cette égalité à gauche par  $Q^T$ , on a

$$Q^T A = Q^T (QR) = (Q^T Q)R = I_n R = R.$$

On en déduit bien

$$R = Q^T A.$$

De plus, pour chaque  $j$ , on a

$$r_{jj} = \vec{q}_j \cdot \vec{a}_j = \|\vec{u}_j\| > 0,$$

#### Unicité de la factorisation.

Supposons  $A = Q_1 R_1 = Q_2 R_2$  deux factorisations QR. Alors  $Q_2^T Q_1 R_1 = R_2$ . En posant  $S = Q_2^T Q_1$  matrice orthogonale  $n \times n$ , on a  $R_2 = S R_1$ , donc  $S = R_2 R_1^{-1}$ . Or  $R_1$  et  $R_2$  sont triangulaires supérieures, donc  $S$  aussi (produit de triangulaires supérieures). Une matrice orthogonale triangulaire supérieure est forcément diagonale avec des entrées diagonales égales à  $\pm 1$ . La positivité des diagonales de  $R_1, R_2$  impose alors  $S = \mathbb{I}_n$ , donc  $Q_1 = Q_2$  et  $R_1 = R_2$ .  $\square$

**Remarques 6.6.0.50.** La condition  $\text{rang}(A) = n$  est équivalente à dire que les colonnes de  $A$  sont linéairement indépendantes, ce qui garantit que tous les vecteurs  $\vec{u}_j$  obtenus par Gram-Schmidt sont non nuls, et donc que  $R$  est inversible avec des coefficients diagonaux strictement positifs.

#### Méthode 6.51

#### Construction de la factorisation QR

1. Appliquer le procédé de Gram-Schmidt aux colonnes de  $A$  pour obtenir une base orthonormée  $(\vec{q}_1, \dots, \vec{q}_n)$
2. Former  $Q = [\vec{q}_1 \ \dots \ \vec{q}_n]$
3. Calculer  $R = Q^T A$  (matrice triangulaire supérieure)

**Exemple.** Soit  $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$ . Trouvons sa factorisation QR.

Les colonnes de  $A$  sont  $\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$  et  $\vec{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ . Elles ne sont pas colinéaires, donc  $\text{rang}(A) = 2$ .

Par Gram-Schmidt :

$$\vec{u}_1 = \vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{q}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{u}_2 = \vec{a}_2 - (\vec{a}_2 \cdot \vec{q}_1) \vec{q}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

$$\vec{q}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \frac{1}{\sqrt{3/2}} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

$$\text{Donc : } Q = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{6} \\ 0 & 2/\sqrt{6} \end{bmatrix}$$

$$\text{Et : } R = Q^T A = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{3/2} \end{bmatrix}$$

**Exemple.** Soit

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}.$$

On cherche une factorisation  $A = QR$ .

Les colonnes de  $A$  sont

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}.$$

**Étape 1 : procédé de Gram-Schmidt.**

*Premier vecteur.*

$$\vec{u}_1 = \vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \|\vec{u}_1\| = \sqrt{1^2 + 1^2} = \sqrt{2}, \quad \vec{q}_1 = \frac{\vec{u}_1}{\|\vec{u}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Deuxième vecteur.

$$\text{proj}_{\vec{u}_1} \vec{a}_2 = \frac{\vec{a}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.$$

On calcule

$$\vec{a}_2 \cdot \vec{u}_1 = 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 + 1 \cdot 0 = 1, \quad \vec{u}_1 \cdot \vec{u}_1 = 2,$$

d'où

$$\text{proj}_{\vec{u}_1} \vec{a}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

On définit

$$\vec{u}_2 = \vec{a}_2 - \text{proj}_{\vec{u}_1} \vec{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

Sa norme vérifie

$$\|\vec{u}_2\|^2 = \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + 1^2 + 0^2 + 1^2 = \frac{5}{2}, \quad \|\vec{u}_2\| = \sqrt{\frac{5}{2}} = \frac{\sqrt{10}}{2},$$

et donc

$$\vec{q}_2 = \frac{\vec{u}_2}{\|\vec{u}_2\|} = \frac{2}{\sqrt{10}} \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \\ 2 \end{bmatrix}.$$

Troisième vecteur.

$$\vec{u}_3 = \vec{a}_3 - \text{proj}_{\vec{u}_1} \vec{a}_3 - \text{proj}_{\vec{u}_2} \vec{a}_3.$$

On calcule d'abord

$$\text{proj}_{\vec{u}_1} \vec{a}_3 = \frac{\vec{a}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1.$$

Comme

$$\vec{a}_3 \cdot \vec{u}_1 = 0 \cdot 1 + 1 \cdot 1 + (-1) \cdot 0 + 2 \cdot 0 + 0 \cdot 0 = 1,$$

on obtient

$$\text{proj}_{\vec{u}_1} \vec{a}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Puis

$$\text{proj}_{\vec{u}_2} \vec{a}_3 = \frac{\vec{a}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \vec{u}_2.$$

On a

$$\vec{a}_3 \cdot \vec{u}_2 = 0 \cdot \frac{1}{2} + 1 \cdot \left(-\frac{1}{2}\right) + (-1) \cdot 1 + 2 \cdot 0 + 0 \cdot 1 = -\frac{3}{2},$$

$$\vec{u}_2 \cdot \vec{u}_2 = \frac{5}{2} \quad \Rightarrow \quad \frac{\vec{a}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} = -\frac{3}{5},$$

et donc

$$\text{proj}_{\vec{u}_2} \vec{a}_3 = -\frac{3}{5} \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/10 \\ 3/10 \\ -3/5 \\ 0 \\ -3/5 \end{bmatrix}.$$

Ainsi

$$\begin{aligned} \vec{u}_3 &= \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} -3/10 \\ 3/10 \\ -3/5 \\ 0 \\ -3/5 \end{bmatrix} \\ &= \begin{bmatrix} -1/5 \\ 1/5 \\ -2/5 \\ 2 \\ 3/5 \end{bmatrix}. \end{aligned}$$

Sa norme vérifie

$$\|\vec{u}_3\|^2 = \left(-\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(-\frac{2}{5}\right)^2 + 2^2 + \left(\frac{3}{5}\right)^2 = \frac{23}{5}, \quad \|\vec{u}_3\| = \sqrt{\frac{23}{5}} = \frac{\sqrt{115}}{5}.$$

Donc

$$\vec{q}_3 = \frac{\vec{u}_3}{\|\vec{u}_3\|} = \frac{5}{\sqrt{115}} \begin{bmatrix} -1/5 \\ 1/5 \\ -2/5 \\ 2 \\ 3/5 \end{bmatrix} = \frac{1}{\sqrt{115}} \begin{bmatrix} -1 \\ 1 \\ -2 \\ 10 \\ 3 \end{bmatrix}.$$

On a ainsi construit une base orthonormée

$$(\vec{q}_1, \vec{q}_2, \vec{q}_3)$$

de  $\text{Im}(A)$ .

**Étape 2 : matrice  $Q$ .**

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{115}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{10}} & \frac{1}{\sqrt{115}} \\ 0 & \frac{2}{\sqrt{10}} & -\frac{2}{\sqrt{115}} \\ 0 & 0 & \frac{3}{\sqrt{115}} \\ 0 & 0 & \frac{3}{\sqrt{115}} \\ 0 & \frac{2}{\sqrt{10}} & \frac{3}{\sqrt{115}} \end{bmatrix}.$$

**Étape 3 : matrice  $R = Q^T A$ .**

Les coefficients de  $R$  sont  $r_{ij} = \vec{q}_i \cdot \vec{a}_j$ . On obtient

$$R = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{10}}{2} & -\frac{3}{\sqrt{10}} \\ 0 & 0 & \frac{\sqrt{115}}{5} \end{bmatrix}.$$

On vérifie que  $A = QR$  et que  $Q$  a des colonnes orthonormées et  $R$  est triangulaire supérieure : c'est une factorisation QR de  $A$ .

## 6.9 La méthode des moindres carrés

La méthode des moindres carrés permet de trouver la « meilleure » solution approximative d'un système d'équations linéaires qui n'a pas de solution exacte.

### Motivation

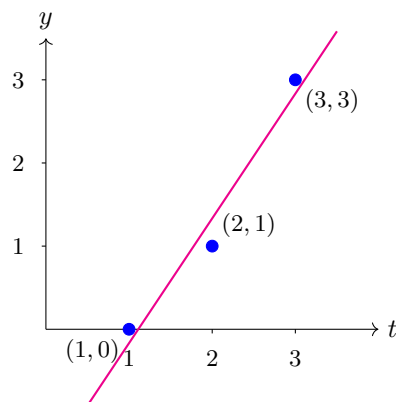
Considérons trois points du plan :

$$(1, 0), (2, 1), (3, 3)$$

On constate qu'ils ne sont pas alignés !

#### Question :

Quelle est l'équation de la droite du plan la plus « proche » de ces points ?



Soit  $y = x_1 + x_2 t$  l'équation cherchée, où  $x_1, x_2$  sont à déterminer.  
On aimerait que les trois points donnés soient sur la droite :

$$(t_1, y_1) = (1, 0) : 0 = x_1 + 1 \cdot x_2$$

$$(t_2, y_2) = (2, 1) : 1 = x_1 + 2 \cdot x_2$$

$$(t_3, y_3) = (3, 3) : 3 = x_1 + 3 \cdot x_2$$

Équation matricielle associée :  $A\vec{x} = \vec{b}$  où  $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$  et  $\vec{b} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$

Matrice augmentée :

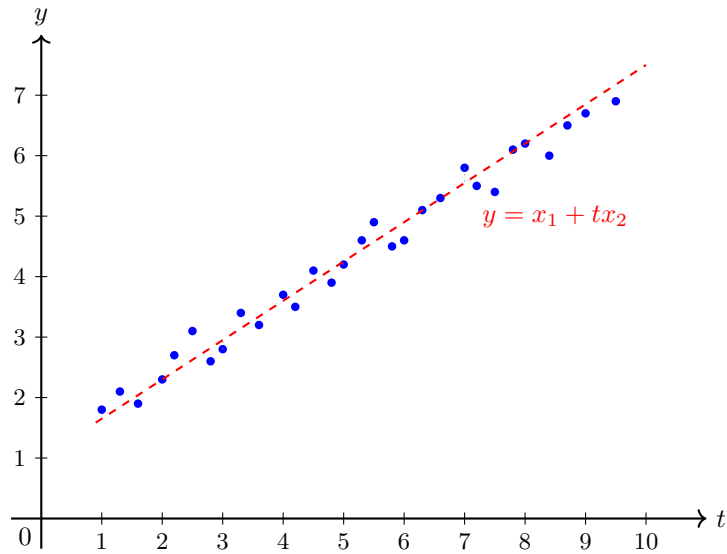
$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Cette équation n'a donc pas de solution (la dernière ligne donne  $0 = 1$ ) ce qui confirme que les points ne sont pas alignés. Aucune droite ne passe donc par ces trois points. Nous allons donc non pas « résoudre » mais le « résoudre approximativement ». C'est un problème très classique de régression linéaire.

### Cas général

Soit  $A$  une matrice de taille  $m \times n$  et  $\vec{b} \in \mathbb{R}^m$ . Nous avons vu que le système  $A\vec{x} = \vec{b}$  possède des solutions si et seulement si  $\vec{b} \in \text{Im}(A)$ . En particulier, si  $\vec{b} \notin \text{Im}(A)$  alors  $A\vec{x} = \vec{b}$  ne possède pas de solution !

En pratique, on s'intéresse à des situations où  $m \gg n$  ( $m$  beaucoup plus grand que  $n$ ). Il est alors, dans les cas pratiques, très probable que  $\vec{b} \notin \text{Im}(A)$ . Dans l'exemple précédent, cela revient à avoir un très grand nombre de points (beaucoup plus que 3), et à chercher une droite qui les relie. Si ces points sont issus de données expérimentales, il est peu probable qu'ils soient alignés. On va alors chercher la droite « la plus proche globalement » de tous ces points.

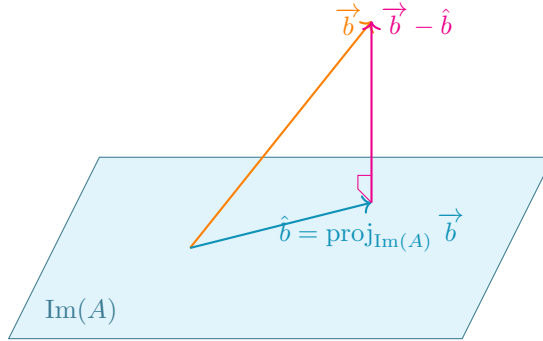


1	1.0	1.8
1	1.3	2.1
1	1.6	1.9
1	2.0	2.3
1	2.2	2.7
1	2.5	3.1
1	2.8	2.6
1	3.0	2.8
1	3.3	3.4
1	3.6	3.2
1	4.0	3.7
1	4.2	3.5
1	4.5	4.1
1	4.8	3.9
1	5.0	4.2
1	5.3	4.6
1	5.5	4.9
1	5.8	4.5
1	6.0	4.6
1	6.3	5.1
1	6.6	5.3
1	7.0	5.8
1	7.2	5.5
1	7.5	5.4
1	7.8	6.1
1	8.0	6.2
1	8.4	6.0
1	8.7	6.5
1	9.0	6.7
1	9.5	6.9

Ici, on cherche à résoudre un système dont la matrice augmentée est  $A =$

Considérons  $\hat{b} = \text{proj}_{\text{Im}(A)} \vec{b}$ .

Géométriquement :



Comme par construction  $\hat{b} \in \text{Im}(A)$ , le système  $A\vec{x} = \hat{b}$  est consistant et il existe au moins un  $\hat{x} \in \mathbb{R}^n$  tel que  $A\hat{x} = \hat{b}$ .

Comme par ailleurs,  $\hat{b}$  est d'après le théorème de la meilleure approximation 6.36, le vecteur de  $\text{Im}(A)$  le plus proche de  $\vec{b}$ , en résolvant  $A\hat{x} = \hat{b}$  au lieu de  $A\vec{x} = \vec{b}$ , on a la meilleure approximation des solutions.

#### Définition 6.52

*Solution des moindres carrés*

Soit  $A$  une matrice  $m \times n$  et  $\vec{b} \in \mathbb{R}^m$ .

Une *solution des moindres carrés* de  $A\vec{x} = \vec{b}$  est un vecteur  $\hat{x} \in \mathbb{R}^n$  qui minimise :

$$\|A\vec{x} - \vec{b}\|$$

Autrement dit :  $\|A\hat{x} - \vec{b}\| \leq \|A\vec{x} - \vec{b}\|$  pour tout  $\vec{x} \in \mathbb{R}^n$ .

La distance  $\|A\hat{x} - \vec{b}\|$  est appelée *erreur de l'approximation*.

Comment calculer  $\hat{x}$  ?

#### Théorème 6.53

$$A\hat{x} = \hat{b} \Leftrightarrow A^T A\hat{x} = A^T \vec{b}$$

Le vecteur  $\hat{x}$  est une solution des moindres carrés de  $A\vec{x} = \vec{b}$  si et seulement si :

$$A^T A\hat{x} = A^T \vec{b}$$

*Démonstration.* ( $\Rightarrow$ ) Commençons par montrer que si  $\hat{x}$  est une solution des moindres carrés de  $A\vec{x} = \vec{b}$  alors  $A^T A\hat{x} = A^T \vec{b}$ .

En notant  $\hat{b} = A\hat{x}$  où  $\hat{x}$  est la solution des moindres carrés, on a  $\vec{b} - \hat{b} \in (\text{Im}(A))^\perp$ .

Donc  $\vec{b} - A\hat{x} \in (\text{Im}(A))^\perp$ , ce qui signifie que  $\vec{b} - A\hat{x}$  est orthogonal à chaque colonne de

$A$ , donc que le produit scalaire entre chaque colonne de  $A$  et  $\vec{b} - A\hat{x}$  est nul.

Par définition du produit scalaire, ceci équivaut à  $A^T(\vec{b} - A\hat{x}) = \vec{0}$ , soit  $A^T A\hat{x} = A^T \vec{b}$ .

( $\Leftarrow$ ) Montrons maintenant qu'une solution de  $A^T A\hat{x} = A^T \vec{b}$  est une solution au sens des moindres carrés de  $A\vec{x} = \vec{b}$ .

Si on a  $A^T A\hat{x} = A^T \vec{b}$ , alors  $A^T(A\hat{x} - \vec{b}) = \vec{0}$ .

Par conséquent,  $\vec{b} - A\hat{x} \in \text{Ker}(A^T) = (\text{Im}(A))^\perp$  (propriété 6.16).

Comme  $\vec{b} = \underbrace{(\vec{b} - A\hat{x})}_{\in (\text{Im}(A))^\perp} + \underbrace{A\hat{x}}_{\in \text{Im}(A)}$ , l'unicité de la décomposition orthogonale nous dit que

$$A\hat{x} = \text{proj}_{\text{Im}(A)} \vec{b} = \hat{b}$$

autrement dit,  $\hat{x}$  est une solution au sens des moindres carrés de  $A\vec{x} = \vec{b}$ .

□

**Remarques 6.6.0.54.** Par construction de  $\hat{b}$  et  $\hat{x}$ , l'équation  $A^T A\hat{x} = A^T \vec{b}$  admet au moins une solution.

#### Définition 6.55

*Système normal*

Soit  $A$  une matrice  $m \times n$ ,  $\vec{b} \in \mathbb{R}^m$ . Le système  $(A^T A)\vec{x} = A^T \vec{b}$  est appelé système normal associé au système  $A\vec{x} = \vec{b}$ .

**Exemples.** Reprenons l'exemple introductif de régression linéaire.

$$\text{Soit } A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \text{ et } \vec{b} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$$

Trouver une solution au sens des moindres carrés du système  $A\vec{x} = \vec{b}$ .

On a

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

et

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 11 \end{bmatrix}$$

Matrice augmentée associée au système normal :

$$\begin{bmatrix} 3 & 6 & 4 \\ 6 & 14 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -5/3 \\ 0 & 1 & 3/2 \end{bmatrix}$$

Donc  $\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} -5/3 \\ 3/2 \end{bmatrix}$  et la droite qui s'approche le mieux (au sens des moindres carrés) des

points  $(1, 0), (2, 1), (3, 3)$  est  $y = -\frac{5}{3} + \frac{3}{2}t$ .

De manière générale, la droite de régression linéaire pour les points

$$(t_1, y_1), (t_2, y_2), \dots, (t_N, y_N) \in \mathbb{R}^2$$

peut être trouvée à l'aide de la solution au sens des moindres carrés du système  $A\vec{x} = \vec{b}$  où

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_N \end{bmatrix} \in \mathcal{M}_{N,2}(\mathbb{R}) \quad \text{et} \quad \vec{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \in \mathbb{R}^N$$

Si  $\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$  est la solution au sens des moindres carrés, alors l'équation de la droite de régression linéaire est :

$$y = \hat{x}_1 + \hat{x}_2 t$$

**Exemples.** Soit  $A = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$  et  $\vec{b} = \begin{bmatrix} 4 \\ 2 \\ 1 \\ 3 \\ 7 \end{bmatrix}$

On constate que le système  $A\vec{x} = \vec{b}$  n'a pas de solution.

Trouver une solution au sens des moindres carrés du système  $A\vec{x} = \vec{b}$ .

On a

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 10 \end{bmatrix}$$

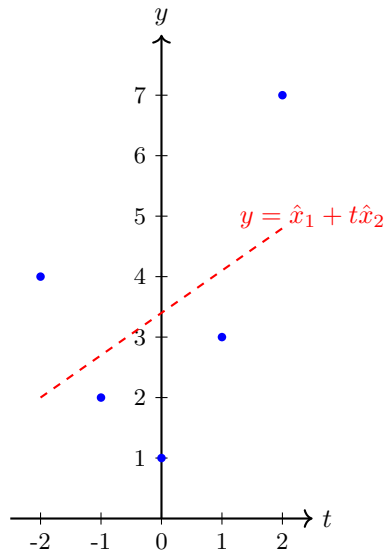
et

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 1 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 17 \\ 7 \end{bmatrix}$$

d'où

$$\begin{bmatrix} 5 & 0 & 17 \\ 0 & 10 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 17/5 \\ 0 & 1 & 7/10 \end{bmatrix}$$

Donc la solution est  $\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} = \begin{bmatrix} 17/5 \\ 7/10 \end{bmatrix}$



**Exemples.** 1. Soit  $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$  et  $\vec{b} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \\ -1 \end{bmatrix}$ .

Trouver une solution au sens des moindres carrés du système  $A\vec{x} = \vec{b}$ .

On a

$$A^T A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 6 \end{bmatrix}.$$

De même,

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

Le système normal est donc

$$(A^T A)\vec{x} = A^T \vec{b}, \quad \text{soit} \quad \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

Matrice augmentée associée :

$$\begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 3 & -1 & 2 \\ 1 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 3/23 \\ 0 & 1 & 0 & 15/23 \\ 0 & 0 & 1 & 2/23 \end{bmatrix}.$$

Donc

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = \begin{bmatrix} 3/23 \\ 15/23 \\ 2/23 \end{bmatrix}.$$

$$2. \text{ Soit } A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix} \text{ et } \vec{b} = \begin{bmatrix} 3 \\ 2 \\ 5 \\ 4 \\ 4 \end{bmatrix}$$

On constate que le système  $A\vec{x} = \vec{b}$  n'a pas de solution.

Trouver une solution au sens des moindres carrés du système  $A\vec{x} = \vec{b}$ .

On a

$$\begin{aligned} A^T A &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 1 & 0 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 4 & 4 \\ 4 & 6 & 1 \\ 4 & 1 & 6 \end{bmatrix} \end{aligned}$$

et

$$A^T \vec{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 1 & 0 & 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 5 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 18 \\ 14 \\ 17 \end{bmatrix}$$

Matrice augmentée associée au système normal :

$$\begin{bmatrix} 5 & 4 & 4 & 18 \\ 4 & 6 & 1 & 14 \\ 4 & 1 & 6 & 17 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & \frac{2}{3} \\ 0 & 1 & 0 & \frac{23}{15} \\ 0 & 0 & 1 & \frac{32}{15} \end{bmatrix}$$

$$\text{Donc } \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 23/15 \\ 32/15 \end{bmatrix}$$

**Théorème 6.56**

Soit  $A$  une matrice de taille  $m \times n$ .

Les affirmations suivantes sont équivalentes :

1.  $A\vec{x} = \vec{b}$  admet une unique solution au sens des moindres carrés pour tout  $\vec{b} \in \mathbb{R}^m$ .
2. Les colonnes de  $A$  sont linéairement indépendantes.
3.  $\text{rang}(A) = n$
4.  $\text{rang}(A^T A) = n$
5. La matrice  $A^T A$  est inversible.

Dans ce cas, la solution au sens des moindres carrés s'écrit

$$\hat{x} = (A^T A)^{-1} A^T \vec{b}$$

*Démonstration.* Les équivalences (1)  $\Leftrightarrow$  (2)  $\Leftrightarrow$  (3) découlent de ce qui précède et du théorème du rang.

L'équivalence (4)  $\Leftrightarrow$  (5) découle du théorème du rang appliqué à  $A^T A$ .

Le seul point à montrer est (3)  $\Leftrightarrow$  (4). Nous allons montrer que  $\text{Ker}(A^T A) = \text{Ker}(A)$  et conclure grâce au théorème du rang.

—  $\text{Ker}(A) \subset \text{Ker}(A^T A)$

Si  $\vec{x} \in \text{Ker}(A)$ , alors  $A\vec{x} = \vec{0}$ , donc

$$A^T A\vec{x} = A^T \vec{0} = \vec{0},$$

donc  $\vec{x} \in \text{Ker}(A^T A)$ .

Ainsi,  $\text{Ker}(A) \subset \text{Ker}(A^T A)$ .

—  $\text{Ker}(A^T A) \subset \text{Ker}(A)$ .

Soit maintenant  $\vec{x} \in \text{Ker}(A^T A)$ , c'est-à-dire

$$A^T A\vec{x} = \vec{0}.$$

On prend le produit scalaire avec  $\vec{x}$  :

$$\vec{x} \cdot (A^T A\vec{x}) = \vec{x} \cdot \vec{0} = 0.$$

En notation matricielle :

$$\vec{x} \cdot (A^T A\vec{x}) = \vec{x}^T (A^T A\vec{x}) = (A\vec{x}) \cdot (A\vec{x}) = \|A\vec{x}\|^2.$$

On obtient donc

$$\|A\vec{x}\|^2 = 0.$$

Par positivité de la norme, cela implique

$$A\vec{x} = \vec{0},$$

c'est-à-dire  $\vec{x} \in \text{Ker}(A)$ .

On a ainsi montré  $\text{Ker}(A^T A) \subset \text{Ker}(A)$ . □

### Théorème 6.57

Soit  $A$  une matrice de taille  $m \times n$  dont les colonnes sont linéairement indépendantes. Soit  $A = QR$  une factorisation QR de  $A$ . La solution au sens des moindres carrés de  $A\vec{x} = \vec{b}$  est unique et s'écrit :

$$\hat{x} = R^{-1}Q^T\vec{b}$$

*Démonstration.* Comme les colonnes de  $A$  sont linéairement indépendantes, la solution au sens des moindres carrés est unique :

$$\hat{x} = (A^T A)^{-1}A^T\vec{b}$$

Par hypothèse,  $A$  peut s'écrire de manière unique sous la forme  $A = QR$  où  $Q$  est une matrice de taille  $m \times n$  telle que  $Q^T Q = \mathbb{I}_n$  et  $R$  est une matrice de taille  $n \times n$  triangulaire supérieure inversible (théorème 6.49).

On a  $A^T = R^T Q^T$  et

$$A^T A = (R^T Q^T)(QR) = R^T(Q^T Q)R = R^T R,$$

d'où

$$(A^T A)^{-1}A^T = (R^T R)^{-1}R^T Q^T = R^{-1}(R^T)^{-1}R^T Q^T = R^{-1}Q^T$$

et  $\hat{x} = R^{-1}Q^T\vec{b}$  □

**Exemple.** Soit  $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$  et soit

$$\vec{b} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}.$$

On a obtenu précédemment une factorisation

$$A = QR$$

avec

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{2}{\sqrt{6}} \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{\frac{3}{2}} \end{bmatrix}.$$

Le théorème assure que la solution des moindres carrés de  $A\vec{x} = \vec{b}$  est

$$\hat{x} = R^{-1}Q^T\vec{b}.$$

Calcul de  $Q^T \vec{b}$ .

$$Q^T \vec{b} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{\sqrt{2}} \\ \frac{3}{\sqrt{6}} \end{bmatrix}.$$

Calcul de  $R^{-1}Q^T \vec{b}$ .

Comme

$$R^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ 0 & \sqrt{\frac{2}{3}} \end{bmatrix},$$

on obtient

$$\hat{x} = R^{-1} \begin{bmatrix} \frac{5}{\sqrt{2}} \\ \frac{3}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} \frac{5}{2} - \frac{3}{6} \\ \frac{2}{2} \\ \frac{3}{3} \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

C'est la solution unique au sens des moindres carrés du système  $A\vec{x} = \vec{b}$ .

**Exemple.** Reprenons la matrice

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 3 \\ 2 \\ -1 \\ 4 \\ 1 \end{bmatrix}.$$

La factorisation QR obtenue précédemment est

$$A = QR$$

avec

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} & -\frac{1}{\sqrt{115}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{115}} \\ 0 & -\frac{\sqrt{10}}{2} & \frac{\sqrt{115}}{2} \\ 0 & \frac{\sqrt{10}}{10} & -\frac{\sqrt{115}}{10} \\ 0 & 0 & \frac{\sqrt{115}}{3} \\ 0 & \frac{2}{\sqrt{10}} & \frac{\sqrt{115}}{5} \end{bmatrix},$$

et

$$R = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{10}}{2} & \frac{3}{\sqrt{10}} \\ 0 & 0 & \frac{\sqrt{115}}{5} \end{bmatrix}.$$

Le théorème dit :

$$\hat{x} = R^{-1}Q^T \vec{b}.$$

**Étape 1 : calcul de  $Q^T \vec{b}$ .**

On obtient

$$Q^T \vec{b} = \begin{bmatrix} \frac{5}{\sqrt{2}} \\ 1 \\ \frac{\sqrt{10}}{44} \\ \sqrt{115} \end{bmatrix}.$$

**Étape 2 : calcul de  $R^{-1}Q^T \vec{b}$ .**

L'inverse de  $R$  est

$$R^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{10}} & -\frac{4}{\sqrt{115}} \\ 0 & \frac{2}{\sqrt{10}} & \frac{\sqrt{115}}{5} \\ 0 & 0 & \sqrt{115} \end{bmatrix}.$$

On obtient donc

$$\hat{x} = R^{-1}Q^T \vec{b} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{10}} & -\frac{4}{\sqrt{115}} \\ 0 & \frac{2}{\sqrt{10}} & \frac{\sqrt{115}}{5} \\ 0 & 0 & \sqrt{115} \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{2}} \\ 1 \\ \frac{\sqrt{10}}{44} \\ \sqrt{115} \end{bmatrix}.$$

Un calcul direct donne

$$\hat{x} = \begin{bmatrix} 20 \\ 23 \\ 31 \\ 23 \\ 44 \\ 23 \end{bmatrix}.$$

C'est la solution unique au sens des moindres carrés du système  $A\vec{x} = \vec{b}$ .

#### Remarque importante 6.58

Nous avons défini  $\hat{b} = \text{proj}_{\text{Im}(A)} \vec{b}$

Comme  $\hat{x} \in \mathbb{R}^n$  est tel que  $A\hat{x} = \hat{b}$ , pour trouver la projection orthogonale de  $\vec{b}$  sur  $\text{Im}(A)$  il suffit de trouver une solution  $\hat{x}$  du système normal  $(A^T A)\vec{x} = A^T \vec{b}$  et la multiplier par  $A$  :

$$\text{proj}_{\text{Im}(A)} \vec{b} = A\hat{x}$$

plutôt que de construire une base orthogonale à l'aide du procédé d'orthogonalisation de Gram-Schmidt.

**Théorème 6.59***Matrice associée à  $\text{proj}_W$* 

Soit  $W$  un sous-espace vectoriel de  $\mathbb{R}^n$  de dimension  $k > 0$ .

Soit  $(\vec{w}_1, \dots, \vec{w}_k)$  une base quelconque de  $W$  (pas forcément orthogonale). Soit  $A$  la matrice de taille  $n \times k$  suivante

$$A = \begin{bmatrix} \vec{w}_1 & \vec{w}_2 & \cdots & \vec{w}_k \end{bmatrix}$$

alors la matrice canoniquement associée à l'application linéaire  $\text{proj}_W : \mathbb{R}^n \rightarrow \mathbb{R}^n$  est

$$A_{\text{proj}_W} = A(A^T A)^{-1} A^T.$$

*Démonstration.* Comme  $(\vec{w}_1, \dots, \vec{w}_k)$  est une base de  $W$  et  $A = \begin{bmatrix} \vec{w}_1 & \vec{w}_2 & \cdots & \vec{w}_k \end{bmatrix}$ , nous avons  $W = \text{Im}(A)$ .

Par conséquent,  $W^\perp = (\text{Im}(A))^\perp = \text{Ker}(A^T)$

Le théorème de la projection orthogonale 6.32 nous dit que tout vecteur  $\vec{v} \in \mathbb{R}^n$  s'écrit de manière unique

$$\vec{v} = \vec{w} + \vec{w}', \text{ où } \vec{w} \in W \text{ et } \vec{w}' \in W^\perp.$$

Comme  $\vec{w} \in W = \text{Im}(A)$ , on a  $\vec{w} = A\vec{x}$ , avec  $\vec{x} \in \mathbb{R}^k$ .

Comme  $\vec{w}' \in W^\perp = \text{Ker}(A^T)$ , on a  $A^T \vec{w}' = \vec{0}$ .

Ainsi  $A^T \vec{v} = A^T(\vec{w} + \vec{w}') = A^T \vec{w} + A^T \vec{w}' = A^T A \vec{x} + \vec{0}$ .

d'où  $A^T \vec{v} = A^T A \vec{x}$ .

Comme  $A^T A$  est une matrice carrée de taille  $k \times k$  et  $\text{rang}(A^T A) = \text{rang}(A)$  (théorème 6.56) la matrice  $A^T A$  est inversible car  $\text{rang}(A^T A) = k$ .

Nous avons donc  $\vec{x} = (A^T A)^{-1} A^T \vec{v}$ , d'où

$$\text{proj}_W \vec{v} = \vec{w} = A\vec{x} = A(A^T A)^{-1} A^T \vec{v}$$

□

**Exemple.** Considérons le sous-espace vectoriel

$$W = \text{Vect} \left( \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right).$$

Une base de  $W$  étant donnée par

$$\vec{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{w}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

la matrice

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

a pour colonnes les vecteurs de cette base.

Le théorème affirme que la matrice associée à la projection orthogonale  $\text{proj}_W : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  est

$$A_{\text{proj}_W} = A(A^T A)^{-1} A^T.$$

Calculons chaque terme.

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

$$(A^T A)^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

Ainsi,

$$A(A^T A)^{-1} = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

Enfin,

$$A_{\text{proj}_W} = A(A^T A)^{-1} A^T = \frac{1}{3} \begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

Ainsi, pour tout  $\vec{x} \in \mathbb{R}^3$  :

$$\text{proj}_W(\vec{x}) = A_{\text{proj}_W} \vec{x} = \frac{1}{3} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix} \vec{x}.$$